心理科学进展 2022, Vol. 30, No. 7, 1604–1611
Advances in Psychological Science

https://doi.org/10.3724/SP.J.1042.2022.01604

# 何以应对道德自我威胁？道德记忆偏差的视角

王修欣 申怡凡

(曲阜师范大学心理学院，曲阜 273165)

**摘 要** 人们会做出一些不道德的事情，这可能对他们积极的道德自我概念构成威胁。为了应对这种道德自我威胁，人们会表现出道德记忆偏差，即遗忘威胁道德自我的不道德事件或信息。近年来，研究者基于自传体记忆范式、游戏范式、代入范式和自我参照范式，为道德记忆偏差提供了证据支持。研究还进一步表明，这种道德记忆偏差可能出于人们应对道德自我威胁的需要。值得注意的是，道德记忆偏差的存在可能需要一定条件。未来研究应该扩展道德记忆偏差的研究范畴，揭示其认知机制，并探究与其他应对道德自我威胁策略的关系。

**关键词** 道德自我，自我保护动机，道德失调，动机性遗忘

**分类号** B848

人们会做出一些不道德的事情，这可能对他们积极的道德自我概念构成威胁(李宏翰，于娟，2013)。那么人们会如何应对这种道德自我威胁呢？已有研究从道德推脱(moral disengagement)和自我合理化(self-serving justifications)等角度回答了该问题(综述见：王珏 等，2016；伍麟，白仲琪，2019)。近期还有系列研究从记忆的角度出发，发现人们会遗忘威胁道德自我的相关事件或信息。这种道德记忆偏差可以使得个体在做出不道德行为后，依然维持良好的道德自我概念。本文聚焦于道德记忆偏差，综述了其研究证据与解释，并讨论了其存在条件和未来研究方向。

## 1 道德记忆偏差现象

### 1.1 自传体记忆范式下的证据

自传体记忆指个体对过去经历或事件的记忆(Rubin, 1986)。自传体记忆范式以被试的自传体记忆为研究对象，通过调查被试对不同性质的自传体记忆的回忆情况来得出结论。比如 Kouchaki 和 Gino (2016)让被试先回忆并写下自己或他人做过的道德、不道德或中性事件，然后评估这些记忆的清晰度和生动性。结果发现，相比于自己做过的道德或中性事件，被试对自己做过的不道德事件的回忆更不清晰、不生动，但是被试对他人做过的不同性质事件的回忆没有显著差别。Ritchie 等(2017)先让被试写下自己或他人做过的许多积极或消极事件(例如：友善地对待他人/背后议论他人)，过段时间再让他们回忆。结果发现，相比于自己做过的消极事件，被试回忆起更多自己做过的积极事件，但是对他人做过的不同性质事件的回忆没有显著差异。

自传体记忆范式以被试的真实经历为研究对象，有着较高的生态效度，但是也存在一些问题。首先，记忆之间的差异不可控。虽然被试回忆的都是道德或不道德事件，但是这些事件在生动性、发生时间和复述次数等方面可能存在差异，这在一定程度上降低了研究的内部效度。其次，研究者没有也往往不能参与到被试的自传体记忆，无法客观地评估记忆的准确性。这使得研究者在对道德记忆偏差进行界定和测量时，只能使用回忆的生动性或频率等指标，而不能使用常见的记忆准确性指标。

### 1.2 游戏范式下的证据

在游戏范式中，研究者通常先让被试完成类似“游戏”的任务，间隔一段时间后，再让其回忆“游戏”中的相关信息。相比于自传体记忆范式，游戏范式创设了研究者可以观察到的真实记忆，使得研究者可以对记忆的准确性进行评估。比如 Shu 和 Gino (2012)让被试完成一个问题解决任务，被

收稿日期：2021-08-26
通信作者：王修欣，E-mail: xxwang@qfnu.edu.cn

试在该任务上的表现越好，获得的奖励就越多。但是被试可以通过谎报自己的表现以获得额外报酬。需要注意的是，研究者在该任务之前，还给被试呈现了一些道德规则，比如不要撒谎等。结果表明相比于没有撒谎的被试，那些撒谎的被试更有可能遗忘这些道德规则。而且在有金钱奖励鼓励被试正确回忆的情况下，这种遗忘依然存在。类似地，Kouchaki 和 Gino (2016)先让被试玩一个抛硬币游戏，被试可以通过撒谎来获得额外收益；两周之后再测量被试对该游戏的记忆，并以被试对实验当天吃晚饭经历的记忆作为对照。结果发现，相比于那些不撒谎的被试，那些撒谎的被试对该游戏的记忆更不生动，也附带更少当时的想法或情绪。而撒谎与不撒谎的被试对吃晚饭经历的回忆没有显著差异。

还有研究者使用独裁者博弈进行了研究。Carlson 等(2020)先让被试扮演分配者角色，决定给接受者多少钱；在完成分心任务后，再让被试回忆给接受者的钱数。结果发现吝啬的即给接受者钱数较少的被试，回忆的给接受者的钱数比实际上给的钱数多，但是慷慨的被试没有表现出该记忆偏差(也见: Tasimi & Johnson, 2015)。当有金钱奖励以鼓励被试正确回忆的时候，这种现象依然存在。Saucet 和 Villeval (2019)则给被试呈现了两种可选的分配方案：自私的分配方案即被试得到的钱数比接受者多，利他的分配方案即被试得到的钱数比接受者少。被试先做出选择，在完成分心任务后，再回忆自己的分配。值得注意的是，在该实验中，研究者会将原来的分配方案呈现出来，但是隐藏了被试所选择的选项中给接受者的钱数。被试的任务是回忆自己选择的选项中给接受者的钱数。结果表明，被试对自私分配方案的回忆正确率显著低于利他分配方案。

游戏范式创设了研究者可以直接观察到的真实记忆，让研究者可以对记忆的准确性进行评估。从这一点来说，游戏范式优于自传体记忆范式。但是在该范式下，研究者根据被试在游戏中的表现，将其区分为道德或不道德的被试。“道德”或“不道德”只是被试的选择，而不是由研究者操纵，这在一定程度上降低了研究的内部效度。

### 1.3　代入范式下的证据

在代入范式中，研究者通常给被试描述一个道德或不道德的行为，让被试以第一或第三人称视角代入。比如 Kouchaki 和 Gino (2016)给被试描述了考试作弊或没有作弊的场景，让被试以第一或第三人称视角代入。在 4 天之后，让被试对这些记忆进行评估。结果表明，在第一人称视角代入时，被试对作弊场景的回忆比没有作弊的场景更不清晰、不生动；但是在以第三人称视角代入时，被试对作弊或没有作弊场景的回忆没有显著差别。

Kouchaki 和 Gino (2016)还让被试代入欺骗或诚实的场景，并在一周后测量了被试对该场景记忆的准确性，发现代入欺骗场景的被试对这些场景的记忆更不准确。Stanley 等(2018)重复了该研究，却发现代入不同场景的被试在记忆的准确性上没有显著差异。而且该结果在采用其它场景，考虑到被试对场景想象程度的差异等因素的可能影响时依然存在。Stanley 等(2018)认为，应该区分记忆的现象学(phenomenology)特征和准确性(accuracy)。现象学特征主要指记忆的生动性和清晰度等。值得注意的是，这些现象学特征与准确性并非完全对应。换句话说，生动或清晰的记忆并不一定是准确的。在代入范式中，道德记忆偏差可能只存在于现象学层面，而不存在于准确性层面。

代入范式虽然有着较高的内部效度，但是尚未得到一致的研究结果，值得未来研究探讨。此外，代入范式是“虚拟的自传体记忆”，而游戏范式是“实验室里的自传体记忆”。换句话说，不管是自传体记忆范式、游戏范式还是代入范式，关注的都是人们对发生在自己身上的事情的情景记忆。那么道德记忆偏差是否也会出现在语义记忆中呢？有研究者使用自我参照范式(the self-reference paradigm)对此进行了研究。

### 1.4　自我参照范式下的证据

自我参照范式通常分为编码和测试两个阶段(刘新明，朱滢，2002)。编码阶段给被试呈现一些词汇，让实验组被试进行自我参照加工(例如：这个词汇能够在多大程度上描述你)，让对照组被试进行他人参照加工(例如：这个词汇能够在多大程度上描述他人)或者语义加工(例如：这个词汇和 XX 的意思相同吗)。在编码阶段后，被试通常先完成分心任务，再进入测试阶段。在测试阶段，除了在编码阶段呈现过的词汇，研究者还会给被试呈现一些新词。被试需要判断这些词汇是否是新词、自我参照、他人参照或者语义参照等。在数

据分析的时候，研究者通常会分析再认回忆和源记忆两个指标。再认回忆主要评估被试能否区分新词和旧词。源记忆主要针对判定为非新词的试次，评估被试能否正确区分词汇是自我、他人或语义参照。

Zhang 等(2018)以积极和消极词汇为编码材料，让被试进行自我和他人参照加工，即让被试判断这些词能否描述他们自己或他人。在这些词汇中，包含许多道德词汇，比如自私-无私等。结果表明，在自我参照条件下，个体对于消极词汇的源记忆明显差于积极词汇；而在他人参照条件下，个体对不同性质词汇的源记忆没有显著差异。而且相比于他人参照，人们在自我参照时对消极词汇的源记忆更差；而人们对积极词汇的源记忆在自我和他人参照时没有显著差异。Rowell 和 Jaswal (2021)给被试呈现友好/礼貌、刻薄/不礼貌和中性的三类跟道德相关的动作词汇(如帮他人开门、推开某人和接住了球)，得到了与 Zhang 等(2018)类似的结果。

这些研究表明，人们更容易忘记或更少记得自身与不道德词汇之间的关联，表现出道德记忆偏差。另外上述研究在源记忆上发现道德记忆偏差，而在再认记忆只发现被试对自我参照词汇的记忆效果好于他人参照。这可能是因为再认记忆更多地反应被试对词汇的加工深度，而源记忆更多地反应被试对词汇之间关系的组织加工(Durbin et al., 2017; Rowell & Jaswal, 2021)。自我参照加工会促进对词汇的深度加工(刘新明，朱滢，2002; Turk et al., 2008)，这可能导致再认记忆时自我参照词汇的记忆效果好于他人参照。自我参照加工也会促进组织加工(刘新明，朱滢，2002; Klein & Loftus，1988)，但是组织加工更容易受到词汇与自我图式的契合性的影响，消极道德词汇可能与自我图式更不契合，从而使得人们消极词汇的源记忆更差(Durbin et al., 2017)。

## 2　道德记忆偏差的解释

目前研究者主要从道德自我威胁的角度对道德记忆偏差给予解释。人们往往有着积极的道德自我概念，而时有发生的不道德事件，可能会对人们的道德自我概念产生威胁；人们可能会为了应对这种道德自我威胁表现出道德记忆偏差。道德记忆偏差主要表现在人们对威胁道德自我的相关事件或信息的遗忘，而人们对于增强道德自我的道德事件或相关信息并没有表现出记忆增强(Rowell & Jaswal, 2021)。这种道德记忆偏差恰好与人们的自我认知相匹配，Klein 和 Epley (2016, 2017)发现人们并不认为自己比他人好，只是没有他人那么坏。这意味着道德记忆偏差主要出于自我保护动机，为应对道德自我威胁的解释提供了间接的证据支持。

从该意义上讲，道德记忆偏差可以在一定程度上汇入记忆忽视效应(the mnemic neglect)。记忆忽视效应指出，人们有着自我保护动机，所以往往会忽视具有自我威胁性的负性信息(综述见 Sedikides et al., 2016)。道德记忆偏差同样基于自我保护动机，只不过主要关注道德自我层面，侧重揭示人们对威胁道德自我的相关事件或信息的选择性遗忘。

道德自我威胁的解释也得到了直接的证据支持。研究发现撒谎的被试对自己有着较低的道德自我评价，而这种道德自我威胁在是否撒谎与道德记忆偏差之间起到中介作用(Kouchaki & Gino, 2016)。道德自我威胁也可以从道德失调的角度来理解，即积极的道德自我与不道德行为之间的失调(Barkan et al., 2015)。研究表明，撒谎被试的确会体验到这种道德失调，而且道德失调在是否撒谎与道德记忆偏差之间起到中介作用(Kouchaki & Gino, 2016)。

研究还表明，在没有道德自我威胁时，道德记忆偏差也不存在。比如这种道德记忆偏差只存在于自我相关的情境，人们对他人相关的事件或信息的记忆不存在差异(Kouchaki & Gino, 2016; Rowell & Jaswal, 2021)。并且这种道德记忆偏差主要出现在做出不道德事件的人身上，比如撒谎的个体(Shu & Gino, 2012)，做出更多自私选择的分配者(Saucet & Villeval, 2019)，或者是违背自身公平原则的分配者(Carlson et al., 2020)。在独裁者博弈中，当分配不是由分配者做出而是随机指定的时候，分配者不需要对分配负责，也就不会体验到道德自我威胁，道德记忆偏差也不复存在(Saucet & Villeval, 2019; Carlson et al., 2020)。

进一步地，道德记忆偏差可能不仅源自人们做出不道德行为后的道德自我威胁，还可能源自人们预期要做不道德行为时的道德自我威胁。这跟 Shalvi 等(2015)所讲的体验到的道德自我威胁

(experienced threat to the moral self)和预期的道德自我威胁(anticipated threat to the moral self)对应。Kouchaki 和 Gino (2016)认为道德记忆偏差可能缓解了预期的道德自我威胁，并引发了后续的不道德行为。Galeotti 等(2020)也指出，道德记忆偏差可以帮助个体实现自我原谅，以合理化未来的不道德行为。

除了上述的动机机制外，研究者还对道德记忆偏差的认知机制进行了探讨。首先，人们表现出道德记忆偏差时，是否真的发生了遗忘？一种可能性是真的发生了遗忘，但是还存在另外一种可能性，即被试明明记得却故意谎报。研究者可以采取某些方式，来减少被试可能的谎报。一种可能的方式是对被试的正确回忆给予奖励。研究者发现即使在有奖励鼓励被试正确回忆的时候，这种记忆偏差依然存在(Carlson et al., 2020; Saucet & Villeval, 2019)。这在一定程度上说明，被试是真的发生了遗忘。

其次，被试真的发生了遗忘，那么其过程机制是怎样的？研究者指出两种可能性：一种是有偏的编码，即个体可能会对做出的道德事件有着更深的编码加工，而对不道德事件有着更浅的编码加工；另一种是抑制提取，即个体可能会主动地抑制对不道德事件或相关信息的提取(Anderson & Hanslmayr, 2014)。Shu 和 Gino (2012)发现这种道德记忆偏差在被试做出不道德行为后会出现，但是在做出不道德行为之前没有。这意味着被试对不同事件的编码加工没有差异。进一步地，他们发现撒谎行为降低了人们对道德词汇的可得性，这意味着道德记忆偏差可能出于人们主动的抑制提取。但是也有研究表明，记忆的动机性遗忘可能发生在编码阶段(比如 Rigney et al., 2021)，这值得未来研究进一步关注。神经机制的研究可为此提供更为可靠的因变量指标，比如研究表明抑制提取过程会削弱与记忆提取伴随的脑电活动(Hu et al., 2015; 2017)，抑制引起的遗忘(suppression-induced forgetting)会增加右背外侧前额叶皮层(dorsolateral prefrontal cortex, dlPFC)和额中回(middle frontal gyrus, MFG)的激活(关旭旭, 王红波, 2021)。

道德记忆偏差还可能发生在储存阶段。记忆的建构观认为，记忆在储存阶段会受到个体经验和心理图式的影响而发生改变(莫映桃，高申春, 2011)。个体往往有着相对积极的道德自我概念(Tappin & Mckay, 2017)。受到这种自我图式的影响，个体可能会对不道德事件或相关信息产生遗忘。记忆的这种建构性可能需要一些时间，当时间越长时，记忆可能有着更多的"操作空间"，从而表现出更大的记忆偏差。Tasimi 和 Johnson (2015)发现，这种道德记忆偏差的确会随着时间的增加而变大。

## 3　道德记忆偏差现象存在的条件

首先，该现象可能只存在于人们有意地做出不道德行为的时候。研究者区分了有意和无意的不道德行为，前者指人们故意做出的、可以意识到的不道德行为，而后者指人们无意间做出的、意识不到的不道德行为(比如内隐偏见等，Gino, 2015)。人们的有意不道德行为往往是为了追逐利益而做出，比如人们会为了获得更多的报酬而有意撒谎(Mazar et al., 2008)或做出不公平的分配(Otto & Bolle, 2015; Rode & Menestrel, 2011)。人们在做出有意不道德行为的时候，应该能够意识到自身行为有违道德标准、体验到道德自我威胁，进而表现出道德记忆偏差。而人们做出无意不道德行为的时候，可能不会体验到道德自我威胁，也就不会表现出道德记忆偏差。值得注意的是，本文中所提及的不道德行为，除非特别提及，都指有意的不道德行为。

其次，该现象可能只存在于人们做出不严重的不道德行为的时候。自我概念维护理论(the theory of self-concept maintenance)指出，人们会为了追逐自身利益做出一定程度的不道德行为；与此同时，人们有一系列心理机制，使得他们可以在做出这些不道德行为的同时，还能维护积极的道德自我概念(Mazar et al., 2008)。这里所说的心理机制，即是人们应对道德自我威胁的方式。值得注意的是，这些不道德行为并非完全不道德，而只是在一定程度上不道德。比如在分配决策中，大部分被试不会把所有的钱都分给自己，而会分给他人一部分(Otto & Bolle, 2015)。这些不太严重的不道德行为，为道德自我威胁的成功应对提供了可能。或许可以将这种不太严重的不道德行为称为应对道德自我威胁的行为策略，而将道德记忆偏差等视为应对道德自我威胁的心理策略。人们可能明晰行为策略和心理策略之间的微妙关系，

并把行为策略限定在心理策略可以应对的范围内。然而当严重违反道德时，这些心理策略可能就“罩不住了”，道德记忆偏差可能就不复存在。现有研究也为此提供了一些证据支持，比如调查发现 46%的暴力犯罪者会体验到侵入性记忆，而且他们对这些犯罪经历的回忆更为详细、生动(Evans et al., 2007a; 2007b)。类似地，人们对严重的违反道德事件的回忆频率反而更高，记忆也更为详细、生动(Huang et al., 2020)。

那么什么样的不道德行为是不严重的？首先，社会规范是判定不道德行为严重程度的重要依据。社会规范传达着特定社会情境对于个体恰当或道德行为的期望，人们也往往据此来对个体行为进行评价(Mcdonald & Crandall, 2015)。不严重的不道德行为应该是没有严重违反社会规范的不道德行为，比如被试为了获得一些经济报酬而谎报自己的表现(Mazar et al., 2008)；而严重的不道德行为则严重违反社会规范，比如上文提及的暴力犯罪。其次，符合道德阈限(moral threshold)可能是对于不严重的具体界定。人们往往并非追求圣人式道德，而秉持好事做一点、坏事也做一点的“常人道德” (甘绍平, 2017)。在此基础上，道德阈限模型进一步指出：人们的行为有着一定的底线，即道德阈限；人们的行为往往不会越过这种道德阈限，而符合道德阈限的行为也使得个体可以维护良好的道德自我概念(Zlatev et al., 2020)。当不道德行为没有越过道德阈限时，个体可以通过道德记忆偏差等心理策略来维护道德自我概念；而一旦越过，道德记忆偏差也就不复存在。

## 4 讨论与展望

现有研究为道德记忆偏差现象及其道德自我威胁机制积累了一些研究证据，但是也存在一些不一致的发现，而且还可能存在其他的机制。未来研究应该注意道德记忆偏差存在的条件，并从研究广度和深度方面，进一步地拓展道德记忆偏差研究。

### 4.1 扩展对道德记忆偏差现象的研究

首先，应该寻求可能的调节变量，以整合一些不一致的发现。比如同样采用独裁者范式，Saucet 和 Villeval (2019)发现被试对利他选择时给接受者钱数的回忆正确率显著高于利己选择，而在给予的钱数上没有表现出偏差；而 Carlson 等(2020)却发现自私的被试记得的给予接受者的钱数显著多于实际给予的钱数。这种结果的不一致，可能是由于两个研究中记忆提取线索不同导致的。Saucet 和 Villeval (2019)给被试呈现自私和利他的两个选项，让被试进行选择；在随后的回忆任务中，会把原来的分配方案呈现出来，这给被试的记忆提取提供了很好的线索。这种线索可能使得被试的记忆偏差变小，即在给予接受者的钱数上没有表现出偏差，而只是对利他选择的回忆正确率更高。但是 Carlson 等(2020)直接让被试回忆给接受者的钱数，这意味着被试在回忆时并没有可以利用的线索，这可能会导致更多的记忆偏差。值得注意的是，记忆提取线索对记忆的影响，可能并非停留于认知层面，还包括动机层面。当存在较多的记忆提取线索时，会给被试记忆的主观建构留下更少的“操作空间”；而在有着较少的记忆提取线索这样的模糊设定下，被试可以充分发挥其主观建构性，表现出更多的动机性遗忘。探讨这些问题，有助于揭示人们使用道德记忆偏差应对道德自我威胁的情境差异。

其次，应探讨道德记忆偏差现象及其泛化的文化差异。目前研究者普遍认为，这种道德记忆偏差主要源自人们应对道德自我威胁的自我保护动机。而跨文化研究表明，东方文化下的个体比西方文化下的个体有着更强的自我保护动机(Sedikides et al., 2015)。那么，这种道德记忆偏差现象是否会在东方文化下的被试群体中表现的更为明显？此外，道德记忆偏差会泛化到他人身上，比如 Ritchie 等(2017)发现，个体会遗忘自己喜欢的人做过的消极事件。但是这种泛化会泛化到哪些他人？从道德自我威胁角度来看，回答这个问题需要搞清楚个体认为自己需要对哪些他人的不道德行为负责，即个体会把哪些他人的行为纳入自身的道德自我概念中。在不同的文化下，人们对将他人纳入自我概念的程度也存在差异，比如 Zhu 等(2007)发现，中国人会把母亲纳入自我概念，而西方人则不会。探讨这些问题，有助于揭示人们使用道德记忆偏差应对道德自我威胁的文化差异。

再者，应进一步扩展道德记忆偏差的研究范畴。目前研究主要关注撒谎和分配领域，而不道德行为还存在于伤害和忠诚等其他领域(Hofmann et al., 2014)。比如在道德困境中，人们可能会为了

救更多的人而伤害少数人(刘传军，廖江群，2021)。在利益冲突的建议给出情境中，建议者可能出于私利而给出有偏的建议(Barneron & Yaniv, 2020)。在这些领域中是否也存在道德记忆偏差现象？而且在道德的不同领域里，道德记忆偏差是否会存在程度上的差异？探讨这些问题，有助于更好地回答人们如何使用道德记忆偏差来应对道德自我威胁。

### 4.2　深入探讨道德记忆偏差的机制

研究者多将道德自我威胁视为一种内部动机，即认为人们想维护在自己眼中的道德形象(Mazar et al., 2008)。类似地，也有研究者将道德记忆偏差称为“自我印象管理的策略” (self-impression management strategy)，即人们不想让自己觉得自己不道德(Saucet & Villeval, 2019)。然而人们不仅想维护在自己眼中的形象，还想维护在他人眼中的形象(Dana et al., 2007)。这种维护在他人眼中形象的印象管理动机，也会影响人们的心理或行为表现，比如人们会为了表现出有能力/热情的形象，而在道德困境中做出更多的符合功利主义/道义论的选择(Rom & Conway, 2018)。人们可能不仅为了应对道德自我威胁，也是为了维护在他人面前的道德形象，而表现出道德记忆偏差。虽然很难将这两者完全区分，但是研究者可以尝试使用公开/私密的情境设置，来探讨其相对作用。

### 4.3　探讨道德记忆偏差与其他应对道德自我威胁策略的关系

从某种程度上讲，道德记忆偏差是一种功能取向的记忆研究(李荆广，郭秀艳，2009)，侧重关注其在应对道德自我威胁中所发挥的作用。值得注意的是，在道德记忆偏差之外，人们还有其他许多应对道德自我威胁的策略，比如前文提及的道德推脱和自我合理化。那么道德记忆偏差与这些策略之间是什么关系呢？这是一个很难回答却也很有价值的问题。一种可能性是，道德记忆偏差与这些策略之间是相关补充、共同发挥作用的。比如Stanley和Brigard (2019)认为，道德记忆偏差和心理远离是互补的：人们会有选择性地遗忘威胁道德自我的不道德事件；而当这些不道德事件没有或者不能被遗忘时，人们可能会在心理上远离，即认为自从那些不道德事件发生后，自己已经发生了很大的改变。还有一种可能性是，其他策略可能为道德记忆偏差的出现提供了条件。人们在做出不道德行为后，可能会通过道德推脱和自我合理化等方式来降低自身的道德标准，改观自己对该事件的评价，从而更容易地模糊不道德行为，表现出道德记忆偏差。这值得未来研究进一步关注，以更加全面地回答人们如何应对道德自我威胁。

**致谢：**感谢匿名评审专家对本文提出的建设性意见！感谢马斯特里赫特大学的张义康博士和华东师范大学的孙倩博士对英文摘要的把关！感谢葛姝纯和李方对本文的挑剔性阅读！

## 参考文献

甘绍平. (2017). 常人道德的尺度. *道德与文明*, (3), 30–36.
关旭旭，王红波. (2021). 抑制引起的遗忘及其神经机制. *心理科学进展*, *29*(4), 665–676.
李宏翰，于娟. (2013). 道德自我受胁及其应对机制. *广西师范大学学报(哲学社会科学版)*, *49*(1), 147–153.
李荆广，郭秀艳. (2009). 记忆研究的功能取向. *心理科学进展*, *17*(5), 923–930.
刘传军，廖江群. (2021). 道德困境研究的范式沿革及其理论价值. *心理科学进展*, *29*(8), 1508–1520.
刘新明，朱滢. (2002). 记忆的自我参照效应. *心理科学进展*, *10*(2), 121–126.
莫映桃，高申春. (2011). 记忆的社会建构论阐释. *心理研究*, *4*(3), 21–27.
王珏，吴明证，孙晓玲. (2016). 道德失调的自我调节策略. *心理科学*, *39*(6), 1473–1478.
伍麟，白仲琪. (2019). 为何不道德行为容易重复发生?——基于心理学的解释. *华中师范大学学报(人文社会科学版)*, *58*(2), 184–192.
Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences*, *18*(6), 279–292.
Barkan, R., Ayal, S., & Dan, A. (2015). Ethical dissonance, justifications, and moral behavior. *Current Opinion in Psychology*, *6*, 157–161.
Barneron, M., & Yaniv, I. (2020). Advice-giving under conflict of interest: Context enhances self-serving behavior. *Journal of Experimental Social Psychology*, *91*, 104046.
Carlson, R. W., Maréchal, M. A., Oud, B., Fehr, E., & Crockett, M. J. (2020). Motivated misremembering of selfish decisions. *Nature Communications*, *11*, 2100.
Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1),

67–80.

Durbin, K. A., Mitchell, K. J., & Johnson, M. K. (2017). Source memory that encoding was self-referential: The influence of stimulus characteristics. *Memory, 25*(9), 1191–1200.

Evans, C., Ehlers, A., Mezey, G., & Clark, D. M. (2007a). Intrusive memories and ruminations related to violent crime among young offenders: Phenomenological characteristics. *Journal of Traumatic Stress, 20*(2), 183–196.

Evans, C., Ehlers, A., Mezey, G., & Clark, D. M. (2007b). Intrusive memories in perpetrators of violent crime: Emotions and cognitions. *Journal of Consulting and Clinical Psychology, 75*(1), 134–144.

Galeotti, F., Saucet, C., & Villeval, M. C. (2020). Unethical amnesia responds more to instrumental than to hedonic motives. *Proceedings of the National Academy of Sciences, 117*(41), 25423–25428.

Gino, F. (2015). Understanding ordinary unethical behavior: Why people who value morality act immorally. *Current Opinion in Behavioral Sciences, 3*, 107–111.

Hu, X., Bergström, Z. M., Bodenhausen, G. V., & Rosenfeld, J. P. (2015). Suppressing unwanted autobiographical memories reduces their automatic influences: Evidence from electrophysiology and an implicit autobiographical memory test. *Psychological Science, 26*(7), 1098–1106.

Hu, X., Bergström, Z. M., Gagnepain, P., & Anderson, M. C. (2017). Suppressing unwanted memories reduces their unintended influences. *Current Directions in Psychological Science, 26*(2), 197–206.

Huang, S., Stanley, M. L., & de Brigard, F. (2020). The phenomenology of remembering our moral transgressions. *Memory and Cognition, 48*(2), 277–286.

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science, 345*(6202), 1340–1343.

Klein, N., & Epley, N. (2016). Maybe holier, but definitely less evil, than you: Bounded self-righteousness in social judgment. *Journal of Personality and Social Psychology, 110*(5), 660–674.

Klein, N., & Epley, N. (2017). Less evil than you: Bounded self-righteousness in character inferences, emotional reactions, and behavioral extremes. *Personality and Social Psychology Bulletin, 43*(8), 1202–1212.

Klein, S. B., & Loftus, J. (1988). The nature of self-referent encoding: The contributions of elaborative and organizational processes. *Journal of Personality and Social Psychology, 55*(1), 5–11.

Kouchaki, M., & Gino, F. (2016). Memories of unethical actions become obfuscated over time. *Proceedings of the National Academy of Sciences, 113*(22), 6166–6171.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*(6), 633–644.

Mcdonald, R. I., & Crandall, C. S. (2015). Social norms and social influence. *Current Opinion in Behavioral Sciences, 3*, 147–151.

Otto, P. E., & Bolle, F. (2015). Exploiting one’s power with a guilty conscience: An experimental investigation of self-serving biases. *Journal of Economic Psychology, 51*, 79–89.

Rigney, A. E., Schnyer, D. M., Hu, X., & Beer, J. S. (2021). Mechanisms of a spotless self-image: Navigating negative, self-relevant feedback. *Self and Identity, 20*(8), 1057–1076.

Ritchie, T. D., Sedikides, C., & Skowronski, J. J. (2017). Does a person selectively recall the good or the bad from their personal past? It depends on the recall target and the person's favourability of self-views. *Memory, 25*(8), 934–944.

Rode, J., & Menestrel, M. L. (2011). The influence of decision power on distributive fairness. *Journal of Economic Behavior and Organization, 79*(3), 246–255.

Rom, S. C., & Conway, P. (2018). The strategic moral self: Self-presentation shapes moral dilemma judgments. *Journal of Experimental Social Psychology, 74*, 24–37.

Rowell, S. F., & Jaswal, V. K. (2021). I remember being nice: Self-enhancement memory bias in middle childhood. *Memory, 29*(2), 261–269.

Rubin, D. C. (1986). *Autobiographical memory*. New York, NY: Cambridge University Press.

Saucet, C., & Villeval, M. C. (2019). Motivated memory in dictator games. *Games and Economic Behavior, 117*, 250–279.

Sedikides, C., Gaertner, L., & Cai, H. (2015). On the panculturality of self-enhancement and self-protection motivation: The case for the universality of self-esteem. *Advances in Motivation Science, 2*, 185–241.

Sedikides, C., Green, J. D., Saunders, J., Skowronski, J. J., & Zengel, B. (2016). Mnemic neglect: Selective amnesia of one’s faults. *European Review of Social Psychology, 27*(1), 1–62.

Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-serving justifications: Doing wrong and feeling moral. *Current Directions in Psychological Science, 24*(2), 125–130.

Shu, L. L., & Gino, F. (2012). Sweeping dishonesty under the rug: How unethical actions lead to forgetting of moral rules. *Journal of Personality and Social Psychology, 102*(6), 1164–1177.

Stanley, M. L., & de Brigard, F. (2019). Moral memories and the belief in the good self. *Current Directions in Psychological Science, 28*(4), 387–391.

Stanley, M. L., Yang, B. W., & de Brigard, F. (2018). No

evidence for unethical amnesia for imagined actions: A failed replication and extension. *Memory and Cognition, 46*(5), 787–795.

Tappin, B. M., & McKay, R. T. (2017). The illusion of moral superiority. *Social Psychological and Personality Science, 8*(6), 623–631.

Tasimi, A., & Johnson, M. K. (2015). A self-serving bias in children's memories? *Journal of Experimental Psychology: General, 144*(3), 528–533.

Turk, D. J., Cunningham, S. J., & Macrae, C. N. (2008). Self-memory biases in explicit and incidental encoding of trait adjectives. *Consciousness and Cognition, 17*(3), 1040–1045.

Zhang, Y., Pan, Z., Li, K., & Guo, Y. (2018). Self-serving bias in memories: Selectively forgetting the connection between negative information and the self. *Experimental Psychology, 65*(4), 236–244.

Zhu, Y., Zhang, L., Fan, J., & Han, S. H. (2007). Neural basis of cultural influence on self-representation. *NeuroImage, 34*(3), 1310–1316.

Zlatev, J. J., Kupor, D. M., Laurin, K., & Miller, D. T. (2020). Being "Good" or "Good Enough": Prosocial Risk and the Structure of Moral Self-Regard. *Journal of Personality and Social Psychology, 118*(2), 242–253.

## How to cope with the threat to moral self? The perspectives of memory bias in moral contexts

WANG Xiuxin, SHEN Yifan

(*School of Psychology, Qufu Normal University, Qufu* 273165, *China*)

**Abstract:** People sometimes behave unethically, which may threaten their self-concept of being moral. To cope with the threat to the moral self, people would forget these past unethical actions or related information more easily. Recent research using autobiographical memory paradigm, game paradigm, take-in paradigm, and self-reference memory paradigm provides evidence for this memory bias. Moreover, research suggest that this memory bias results from people's need to cope with the threat to their moral self. That is, people selectively forget their unethical behaviors to maintain a positive moral self. Notably, it is only under certain conditions that this memory bias occurs. Future research should provide much more convergent evidence for the phenomenon, examine its underlying neural and cognitive mechanisms, and explore its relationships with other strategies that people use to cope with the threat to their moral self.

**Key words:** moral self, self-protective motivation, ethical dissonance, motivated forgetting